RT-PCR kit, Qiagen GmbH, Germany) according to the manufacturer's instructions. An SP6 VH FR4 specific primer 183187 (5' TCC CTT GGC CCC AGT AAG 3', positions 1247-1264 in GeneBank Accession No. X56936), and a primer specific to the 3' part of SP6 Vκ leader, 183038 (5' CGC GTT TCT CTG GTT GTC 3', positions 224-241 in GeneBank Accession No. J00569) were used, in an annealing temperature of 48°C. The results are shown in **Fig. 8.** The expected product of 743 bp, comprising the 3' end of SP6 Vκ leader, whole V-Jκ, the linker and practically all SP6 VH, was detected only in transfections 4, 5 and 6.

Example 6

Single and double TS events detected in the COS7 transfections

RT-PCR products of Example 5 were either cloned directly (for transfections 5 and 6) or following a nested PCR (for clone 4) into pGEMT. The DNA sequence of five independent clones from each transformation was determined. Five out of five clones generated from both transfections 4 and 5 (**Figs. 8A** and **8B**, respectively), showed the expected sTS joining product: The linker coding sequence was spliced in-frame to the respective SP6 V exon, utilizing the corresponding splice sites

Two very close but distinguishable RT-PCR products were observed for transfection 6 and were co-eluted from an agarose gel for cloning into pGEMT and DNA sequencing. Three out of the five clones generated, all corresponding to the lower molecular weight fragment of the two, showed accurate joining of Vκ to VH, involving the Vκ donor and VH acceptor splice sites (**Fig. 8C**). However, the peptide linker exon was absent from these products, nor could it be identified in a larger number of clones screened for the presence of the linker BamHI site. Thus, this RT-PCR product bears the exact sequence expected from fTS. The other two

clones had a longer DNA insert and contained an irrelevant stretch between Vκ and VH.

Example 7

DNA constructs designed for antibody library construction and their modes of application

Faithful reconstitution of an antibody immune repertoire via dTS, as described herein, requires that all or, at least, the vast majority of the functional antibody H and L chain transcripts, are accessible as substrates of dTS. In other words, for each organism chosen for library production (human, mouse or others), the DNA construct assembled for this purpose should direct the synthesis of dTS-driving transcripts which contain multiple sequences. The sense strand of one part of the construct should harbor sequences reverse complementary (hence, capable of hybridizing) to most of the organism's H chain transcripts, and the other strand should include sequences which can hybridize to the organism's VL transcripts.

According to the specific example presented above, the 5' part of the dTS construct should contain sequences corresponding to the different JL-CL introns, and, more specifically, to the 3' end of the intron including the BP site and the acceptor splice site. In mice, there is only one functional Cκ gene, responsible for more than 95% of mouse antibody light chains, and 4 Cλ genes (hence, 4 relevant introns). In humans, there is also one Cκ gene, responsible for approximately 60% of the human L chains, and 4-5 functional Cλ genes (and J-C introns).

Ideally, the dTS construct should contain sequences with sufficient reverse-complementarity to all these L chain gene introns in a given organism. The 3' part of the construct should harbor sequences which can hybridize upstream to exons encoding all the functional VHs of the organism. There are 39 such genes in humans and approximately the same number in mice. This part of the construct should, therefore, include an array of sequences, each corresponding to one or more VH genes, so that together they correspond to all the functionally expressed VH genes.

In the reciprocal embodiment of the approach (resulting in a library of scFv genes in which VH precedes VL), the 5' part of the construct should possess sequences corresponding to all intronic sequences 5' to all the H chain constant region genes in a certain organism. In humans, these will be Cμ, Cγ1-4, Cα1-2 and Cε, while Cδ can be excluded. Similarly to the previous situation, the 3' part of the construct should include sequences corresponding to all the functional VL genes.

The first design, namely, the one in which VL precedes VH, is simpler, because it requires the incorporation of a smaller number of sequences. We have, therefore, chosen this configuration for the constructs designed for the production of both mouse and human libraries as described below.

The products of accurate fTS or dTS carried out in a large population of B cell clones in the immunized human or animal should constitute a large collection (pool) of transcripts encoding scFv joined to antibody H chain constant region genes. If the first design is employed, then Vκ is upstream to VH, and the inverse is true for the second design. The cloning steps which are required for the production of a phage-display (or any other genetic-display) antibody library from these cells at this stage according to the present invention are, in principle, similar to those which are used in the production of current antibody libraries. However, a major difference, which should render the generation of TS-mediated libraries much simpler and considerably more attractive and efficient, is the need to clone only one gene into each vector molecule, encoding the entire scFv (including the peptide linker), instead of two (VH and VL), in current technologies.

Several display technologies using bacteriophage, yeast, ribosomes or other vehicles for the production of immune antibody libraries have been described and are in current use. Even for a single vehicle, such as bacteriophage, quite a few procedures exist, which differ mainly in finer details. One of the hallmarks of TS-mediated immune antibody production according to the invention is its intrinsic adaptability to practically all display schemes and technologies.

Production of a dTS-based immune antibody library can be carried out by well-known methods published in the literature such as the protocols and reagents

described in (Barbas III et al., 2001) (hereinafter "the Manual"). The Manual is herein incorporated by reference as if fully described herein in its entirety.

In the present example, production of a scFv library from an immunized mouse transgenic for the dTS construct is described based on protocols and reagents described in the Manual. Whenever appropriate, the relevant chapter of the Manual will be cited.

According to this procedure, the DNA construct of the invention designed to mediate fTS or dTS of Ig H and L chain mRNAs is introduced as a transgene into the genome of mice used for immunization and production of mAbs. Trans-spliced mRNA species are expected to be found in all B cells of the transgenic animals.

Immunization schedules, immune serum titering by ELISA, isolation of total RNA from immune cells and first strand cDNA synthesis from total RNA using an oligo (dT) primer and RT, are all standard procedures independent of the scheme of the present invention, and are carried out as described in chapter 8 of the Manual.

As described in Protocol 9.5 of the Manual, for construction of mouse scFv libraries by overlap extension, with either a short (7 amino acids) or a long (18 amino acids) peptide linker, in which VL is placed before VH, two rounds of PCR are required. In the first round, cDNA prepared from total RNA is amplified in three separate reactions. For the Vκ reaction, a mixture of 17 sense primers (MSCVK1-17), corresponding to the 5' end of mouse Vκ genes and combined in specified ratios, is applied together with a mixture of 3 reverse primers, corresponding to mouse Jκ segments (also in specified ratios). There is one set of Vκ reverse primers for a short linker library and one set for a long linker. For Vλ amplification reaction, one sense primer (MSCVL-1) is applied together with one reverse primer (one for short and one for long linker). For the VH reaction, a mixture of 19 sense primers in specified ratios (MSCVH1-19), corresponding to mouse VH genes, is applied together with a mixture of 3 reverse primers specific for the 5' end of the first mouse constant regions (MSCG1ab-B, MSCG3-B, MSCM-B). The Vκ and Vλ reverse primers on one hand and the VH sense primers, on the other hand, are designed to harbor sufficient reverse complementarity to

each other so that one strand of the Vκ or Vλ PCR product can serve, in the second round of PCR, as a primer specific for the other strand of the VH PCR product (overlap extension). The peptide linker is encoded by sequences incorporated into the Vκ or Vλ reverse primers and the VH sense primers. The resulting PCR products of this first round are purified from agarose gels and mixed in correct ratios for the second round of PCR, with the addition of two overlap extension primers: RSC-F (sense) and RSC-B (reverse).

The scFv-encoding products obtained in the second PCR round are purified from agarose gels, cut with SfiI restriction enzyme and cloned into either the pComb3HSS or pComb3XSS phagemid vectors. SfiI recognizes the sequence GGCCNNNNNGGCC, where N can be any one of the 4 nucleotides, and cuts between the 4^{th} and the 5^{th} N. The freedom in N nucleotide selection enables the design of different sites recognized and cut by the same restriction enzyme, hence, the positional cloning mediated by a single enzyme. One such SfiI site is incorporated into all the sense primers specific for VL, and a different SfiI site is incorporated into the VH-specific reverse primers. These sites are correctly positioned in the cloning vectors, so that resulting inserts can be expressed as scFv fusion proteins on the surface of the M13 bacteriophage.

Using primers, vectors and the SfiI cloning design described above, construction of a dTS-mediated mouse scFv library of the invention requires only one round of PCR. For example, for Vκ-VH scFvs, the 17 Vκ sense primers (MSCVK1-17), mixed in the correct ratios, are applied together with the proper mixture of the 3 VH reverse primers (MSCG1ab-B, MSCG3-B, MSCM-B. It should be noted that, in practice, since no second round of PCR is required, the SfiI sites in the PCR products may prove too close to the ends of the DNA molecules, which may require synthesis of longer primers. After SfiI cleavage, the resulting PCR products will only be distinguishable from the products of the second PCR instructed by the Manual by the actual sequence which encodes the peptide linker. These SfiI-cleaved single round PCR products are similarly purified from agarose gels and inserted into either the pComb3HSS or pComb3XSS phagemid vectors. All further steps in library construction, selection of binding products from the

library and analysis of selected antibodies, can then be performed as described in chapters 9-11 of the Manual.

Example 8

DNA constructs designed for TCR library construction and their modes of application

The assembly of TCR dTS constructs is performed as described for antibodies and they comprise the following elements:

1. A fragment with a sense strand reverse complementary to the 3' end of the intron upstream to Cα, including the BP and acceptor splice site.
2. A spacer sequence.
3. A fragment encoding the peptide linker as an exon, preceded by a BP and an acceptor splice site and followed by a donor splice site.
4. A spacer sequence.
5. A fragment comprising a series of stretches, each with a sense strand reverse complementary to one or more TCR Vβ transcripts, in a region centering around the leader donor splice site.

In the reciprocal scenario, element 1 possesses stretches reverse complementary to the 3' end of the introns upstream of both Cβ (in humans there are two functional Cβ genes), including the BP and acceptor splice site. The stretches in element 5 each is reverse complementary to one or more TCR Vα genes, centering around the leader donor splice site.

In fact, elements 2, 3 and 4 may be incorporated into all dTS constructs, whether for antibody or TCR libraries, from any vertebrate species chosen. Of course, the optimal amino acid sequence of the peptide linker for generating functional TCR scFv may be different than that for antibodies, in which case it should replace the antibody linker.

TCR library construction does not differ in essence from that of an antibody library. **Fig. 18** is a schematic representation of the dTS gene of the invention for

generating TCR scFv libraries, and the predicted scFv-encoding RNA product attached to TCR β chain constant region gene.

The construction of a human immune TCR scFv library, in which Vα precedes Vβ, is carried out based on the procedures described in the Manual. Total RNA, extracted from T cells of an immune individual, serves as a template for cDNA synthesis using an oligo (dT) primer. The single round of PCR is performed with an adequate mixture of sense primers specific for the 5' end of all human Vα genes, and a more limited set of reverse primers specific for all human Jβ segments. A specific SfI site is incorporated in all sense primers and another SfiI site is present in all reverse primers. The SfiI-cleaved PCR products are directly subjected to positional cloning into an appropriate phagemid vector specifically designed for TCR scFv expression on phage.

The complex nature of TCR ligand and the much lower affinity of interactions compared with antibodies, may require screening of the library and isolation of clones with multivalent MHC-peptide antigens, and this can be accomplished, for example, by the use of MHC-peptide tetramers (McMichael and O'Callaghan, 1998). These are tetrameric structures in which four biotin-harboring single-chain MHC-peptide proteins are linked by one avidin or streptavidin molecule.

Example 9

Construction of mouse library dTS gene

In order to be able to mediate dTS to all the mouse VH repertoire, we constructed a synthetic DNA fragment, which is analogous to the XbaI/EcoRI fragment in the SP6 dTS mediating construct (1034-2 – see Example 3e above, Fig. 7). This fragment comprises 19 sequences of approx. 20 bases each, so that the RNA stretch encoded by each of these sequences is reverse complementary to the 3' end of one or more mouse VH leader exons and the 5' end of the following intron, including the donor splice site.

These sequences are based on a comparison of the leader exons and flanking sequences of 42 functional mouse VH genes found in the ABG (Germline gene

directories of mouse: http://www.ibt.unam.mx/vir/V_mice.html). In order to facilitate the cloning procedure of the synthetic DNA fragment by complementation, we included several restriction sites in-between the 19 sequences, and an XbaI and an EcoRI sites in its 5' and 3' ends, respectively. The 19 sequences are summarized in **Fig. 9**. **Fig. 11** shows the sequence of the double stranded synthetic DNA fragment encoding these 19 sequences. This 420 bp double-stranded DNA fragment was produced from 11 synthetic oligonucleotides (shown in **Fig. 10**), following single step annealing and ligation reactions. The entire fragment was isolated from an agarose gel, and subcloned for DNA sequence verification. The correct XbaI/EcoRI fragment replaced the SP6 VH specific sequence in clone 1034-2 , producing clone 1068-2 (**Fig.12A**). Complete nucleotide sequence of this clone is shown in **Fig. 12 B**.

Example 10

Construction of human library dTS gene

10a. The human VH targeting sequence

Similarly to the synthetic fragment described in the generation of the mouse dTS clone 1068-2 in Example 9 above, we have designed a synthetic DNA fragment specific for all functional human VH genes. This fragment comprises 24 sequences of approx. 20 bases each, so that the sense strand of each sequence is reverse-complementary to the 3' end of one or more human VH leader exons, and the 5' end of the following intron, including the donor splice site. These sequences are based on a comparison of the 39 functional human VH germ-line genes summarized at IMGT, the International ImMunoGeneTics database (http://imgt.cines.fr:8104). These target sequences are shown in **Fig. 13**.

The expected synthetic fragment, flanked by an XbaI and an EcoRI sites and harboring additional restriction sites to facilitate complementation (if required during the cloning procedure), is presented in **Fig. 14**.

10b. The human VL targeting sequence

In order to be able to mediate dTS to all the human VL repertoire, we have designed a synthetic DNA fragment, which is analogous to the 5' XhoI/NotI fragment in the mouse dTS mediating constructs (1034-2 and 1068-2, see Example 9 above). This new fragment comprises three sequences which encode RNA stretches reverse complementary to the 3' end of all human JL-CL introns as specified below, including the predicted BPs and acceptor splice sites:

1. A 120 b sequence at the 3' end of the human Jκ-Cκ intron. (positions 211- 333 in GenBank Accession No. J00241). The predicted BP is located 83 b upstream to the intron's acceptor (underlined).

5' TCTGGGATAA GCATGCTGTT TTCTGTCTGT CCCTAACATG CCCTGTGATT ATCCGCAAAC AACACACCCA AGGGCAGAAC TTTGTTACTT AAACACCATC CTGTTTGCTT CTTTCCTCAG 3'

2. A 120 b sequence at the 3' end of the human Jλ1-Cλ1 intron (positions 3897-4007 in GenBank Accession X51755). Location of the BP in this region is not clear.

5' CCCCGGGTGG ACCGGATGGC CACACTGTGA ACCCTCCCAG AGACTTTAGA CAGAGAGAGG GGCTCCACAA CACCCCGGTA TTCTGTCTGC CCTCTCTCAC CCCCTTCCCT GTCCACACAG 3'

3. A 50 bp sequence at the 3' end of the human Jλ2-Cλ2 intron (positions 9481-9600 in GenBank Accession X51755); Jλ3-Cλ3 intron (positions 14869-14988 in X51755); Jλ6-Cλ6 intron (positions 28037-28156 in X51755); Jλ7-Cλ7 intron (positions 31081-31200 in X61755). This sequence is identical in all these introns, and it spans the BP which is located 25 bp upstream to the acceptor site.

5' CCCAGGTGGA CACCAGGACT CTGACCCCCT GCCCCTCATC CACCCCGCAG 3'

In order to facilitate the cloning procedure of the synthetic DNA fragment, we included restriction sites in-between the three sequences, and Xho I and Not I sites in its 5' and 3' ends, respectively. **Fig. 16** shows the sequence of the double stranded synthetic DNA fragment that encodes the reverse-complementary sequences. This 316 bp double-stranded DNA fragment is ligated from nine pre-

annealed synthetic oligonucleotides (**Fig. 15**). Following DNA sequence verification, the entire fragment is cloned as an XhoI/NotI fragment in the configuration described for clone 1068-2. The scheme of the human dTS construct is shown in Fig. **17A**, and its complete nucleotide sequence is given in Fig. **17B**.

DISCUSSION

The two sTS constructs examined in our in-cell systems enabled us to dissect the expected dTS process into its separate components. Joining the peptide linker sequence in-frame to the corresponding V gene, as described for clones 965-14, 1033-1 and 1026-2, allowed us to screen for TS RNA products via RT-PCR using exactly the same protocol and DNA primers we useD for the identification of scFv-encoding dTS products.

Our experimental data indicate that both sTS events occur exactly as expected in transfected cells. In COS7 cells, products of the expected size were abundant after one PCR round, and DNA sequencing indicated that 5 out of 5 clones for each of the configurations exhibited the nucleotide sequence expected following TS. Although these results are not quantitative, they are indicative of a considerable frequency of the expected events. In this regard, the observation (see the ELISA results in Example 4), that cells transfected with constructs 1083-1 and 1026-2 secreted approx. 140 ng/ml of human IgG, which could only result from TS, is most informative.

It thus can be concluded that: 1) The splice sites flanking the linker exon, which we have incorporated in the sTS genetic constructs, are recognized and cleaved properly by the spliceosome. 2) The sequences chosen for base-pairing with the relevant regions in the H and L chain transcripts are adequate with respect to both length and target sequences.

Current models propose that splicing normally takes place co-transcriptionally. The probability of TS of transcripts from different chromosomal templates is thus a critical issue, as this will be the setting in antibody-producing cells from which immune libraries are to be generated. In this regard, the SP6 system, in which precise TS occurred between transcripts from the chromosomally-

integrated sTS construct 965-14 and the endogenous H chain RNA in a stable transfectant, lends strong support to our working hypothesis according to the invention. In fact, to the best of our knowledge, this is the first demonstration of experimental base-pairing-directed TS between RNA of two different chromosomal genes in mammalian cells. Although it cannot be ruled out, at this stage, that a fortuitous integration event could have placed the introduced gene in close vicinity in the chromatin to the active SP6 H chain locus, we tend however to interpret this observation as evidence that at least some fraction of the primary transcripts, which are not fully spliced, can still leave the site of transcription and serve as TS substrates at more distal sites in the nucleus.

The nucleus of a COS7 cell transfected with plasmids harboring SV40 origin of replication contains, within 3 days, up to 10(5) extra-chromosomal plasmid copies. Under these conditions, the high concentration of templates and their mode of distribution within the nucleus render a possible limitation by co-transcriptional splicing irrelevant. Hence, this system is ideal for the establishment of important components and parameters, which govern TS efficiency.

In addition to the unambiguous demonstration of the two sTS processes, we could easily detect (after only one round of PCR) a spliced SP6 Vκ-VH mRNA, in which the 3' end of Jκ is precisely joined to the 5' end of the VH exon (see **Fig. 8C**). We deduce from the nucleotide sequence at the junction between the two V genes that joining has occurred through TS. In view of the absence of the peptide linker exon from this junction, we cannot rule out the possibility that these apparent TS products were generated spontaneously, irrelevant of the dTS construct introduced into these cells. However, even when extremely sensitive nested RT-PCR protocols were used, we have never been able to detect such products in SP6 cells or in COS7 controls transfected with SP6 H and L chain genes alone (without the dTS construct). These control transfections were successful and resulted in the production of a high level of IgG (transfections 1 and 2, see ELISA results in Example 4).

Assuming dTS-mediated joining of the SP6 Vκ and VH, the lack of the linker exon raises the possibility of exon skipping, which is one of the mechanisms

responsible for alternative splicing (Smith and Valcarcel, 2000). Since both sTS components of the dTS construct proved functional, it is possible that their specific combination leads to the observed skipping. The inclusion of an exon splicing enhancers [ESE, see (Caceres and Krainer, 1997), which usually appear as typical stretches of purines, may be considered. However, ESEs function is associated with intrinsically weak splicing signals which flank the exon. In the dTS construct, the splice sites perfectly comply with consensus sequences, and they receive high scores when tested with 'Splice Site Prediction by Neural Network' (Berkeley Drosophila Genome Project, http://www.fruitfly.org/seq_tools/splice.html). In addition, a stretch of seven purines which is included in the linker exon, may display a certain degree of enhancer functions. The presence of a splicing silencer in the linker exon cannot be ruled out, as silencer motifs are still obscure. Since alternative splicing is often tissue-specific, it is conceivable that the speculated exon skipping observed in COS7 cells would not normally take place in B cells.

The Vκ-VH sequence cloned from COS7 transfectants encodes extra 4 amino acids, as compared with the mature VH polypeptide. This is an inevitable result of the structure of antibody V genes, where the 5' end of the second exon encodes the C terminus of the leader peptide. The protein products encoded by this joined sequence are likely to result in the formation of diabody-like structures when expressed as dimers. It remains to be seen whether these scFv-encoding genes can indeed be expressed as functional diabodies on the surface of bacteriophages.

These RT-PCR products are identical in sequence to those expected following fTS events. We believe that this is the first demonstration of a directed joining of two different RNA species via TS. In view of the apparent abundance of these products in transfected COS7 cells, we expect fTS to occur in sufficient frequency in normal antibody producing B cells, so that all B cell clones in the body are represented in the resulting library.

REFERENCES

Barbas III, C. F., D. R. Burton, et al. (2001). Phage Display: A Laboratory Manual. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

Bruggemann, M. and M. S. Neuberger (1996). "Strategies for expressing human antibody repertoires in transgenic mice." Immunol. Today **17**(8): 391-7.

Caceres, J. F. and A. R. Krainer (1997). Mammalian pre-mRNA splicing factors. Eukaryotic mRNA Processing. A. R. Krainer. New York, NY, Oxford University Press: 174-212.

Carroll, W. L., E. Mendel, et al. (1988). "Hybridoma fusion cell lines contain an aberrant kappa transcript." Mol. Immunol. **25**(10): 991-5.

Caudevilla, C., D. Serra, et al. (1998). "Natural trans-splicing in carnitine octanoyltransferase pre-mRNAs in rat liver." Proc. Natl. Acad. Sci. U S A **95**(21): 12185-90.

Chiara, M. D. and R. Reed (1995). "A two-step mechanism for 5' and 3' splice-site pairing." Nature **375**(6531): 510-3.

Custodio, N., M. Carmo_Fonseca, et al. (1999). "Inefficient processing impairs release of RNA from the site of transcription." EMBO J. **18**(10): 2855-66.

Daneholt, B. (1999). "Pre-mRNP particles: From gene to nuclear pore." Curr. Biol. **9**(11): R412-5.

Embleton, M. J., G. Gorochov, et al. (1992). "In-cell PCR from mRNA: amplifying and linking the rearranged immunoglobulin heavy and light chain V-genes within single cells." Nucleic Acids Res. **20**(15): 3831-7.

Garcia-Blanco, M. A., M. Puttaraju, et al. (2000). "Spliceosome-mediated RNA trans-splicing in gene therapy and genomics." Gene Ther. Reg. **1**(2): 141-163.

Gluzman, Y. (1981). "SV40-transformed simian cells support the replication of early SV40 mutants." Cell **23**(1): 175-82.

Gross, G. and Z. Eshhar (1992). "Endowing T cells with antibody specificity using chimeric T cell receptors." FASEB J. **6**(15): 3370-8.

Harlow, W. and D. Lane (1988). Antibodies: A laboratory manual, Cold Spring Harbor Laboratory: 139-243.

Holliger, P., T. Prospero, et al. (1993). ""Diabodies": small bivalent and bispecific antibody fragments." Proc. Natl. Acad. Sci. U S A **90**(14): 6444-8.

Huang, X. Y. and D. Hirsh (1992). "RNA trans-splicing." Genet. Eng. (N Y) **14**: 211-29.

Kohler, G. and M. J. Shulman (1980). "Immunoglobulin M mutants." Eur. J. Immunol. **10**: 467-476.

Konarska, M. M., R. A. Padgett, et al. (1985). "Trans splicing of mRNA precursors in vitro." Cell **42**(1): 165-71.

Lin, A. Y., B. Devaux, et al. (1990). "Expression of T cell antigen receptor heterodimers in a lipid-linked form." Science **249**(4969): 677-9.

Lubin, I., Y. Faktorowich, et al. (1991). "Engraftment and development of human T and B cells in mice after bone marrow transplantation." Science **252**(5004): 427-31.

McGuinness, B. T., G. Walter, et al. (1996). "Phage diabody repertoires for selection of large numbers of bispecific antibody fragments." Nat. Biotechnol. **14**(9): 1149-54.

McMichael, A. J. and C. A. O_Callaghan (1998). "A new look at T cells." J. Exp. Med. **187**(9): 1367-71.

Mendez, M. J., L. L. Green, et al. (1997). "Functional transplant of megabase human immunoglobulin loci recapitulates human antibody response in mice." Nat. Genet. **15**(2): 146-56.

Neuberger, M. S. (1983). "Expression and regulation of immunoglobulin heavy chain gene transfected into lymphoid cells." Embo Journal **2**(8): 1373-8.

Nissim, A., H. R. Hoogenboom, et al. (1994). "Antibody fragments from a 'single pot' phage display library as immunochemical reagents." EMBO J. **13**(3): 692-8.

Ochi, A., R. G. Hawley, et al. (1983). "Transfer of a cloned immunoglobulin light-chain gene to mutant hybridoma cells restores specific antibody production." Nature **302**(5906): 340-2.

Puttaraju, M., S. F. Jamison, et al. (1999). "Spliceosome-mediated RNA trans-splicing as a tool for gene therapy." Nat. Biotechnol. **17**(3): 246-52.

Sambrook, J., E. F. Fritsch, et al. (1989). Molecular Cloning - A Laboratory Manual. Cold Spring Harbor, USA, Cold Spring Harbor Laboratory Press.

Shimizu, A., M. C. Nussenzweig, et al. (1989). "Immunoglobulin double-isotype expression by trans-mRNA in a human immunoglobulin transgenic mouse." Proc. Natl. Acad. Sci. U S A **86**(20): 8020-3.

Smith, C. W. and J. Valcarcel (2000). "Alternative pre-mRNA splicing: the logic of combinatorial control." Trends Biochem. Sci. **25**(8): 381-8.

Solnick, D. (1985). "Trans splicing of mRNA precursors." Cell **42**(1): 157-64.

Wang, J., L. G. Cao, et al. (1991). "Localization of pre-messenger RNA at discrete nuclear sites." Proc. Natl. Acad. Sci. U S A **88**(16): 7391-5.

Ward, E. S. (1992). "Secretion of T cell receptor fragments from recombinant Escherichia coli cells." J. Mol. Biol. **224**(4): 885-90.

Watanabe, T. and B. A. Sullenger (2000). "RNA repair: a novel approach to gene therapy." Adv. Drug Deliv. Rev. **44**(2-3): 109-18.

Whitlow, M. and D. Filpula (1992). Single-chain Fvs. Tumor Immunology - A Practical Approach. G. Gallagher, R. C. Rees and C. W. Reynolds. New York, Oxford University Press: 279-291.

Winter, G., A. D. Griffiths, et al. (1994). "Making antibodies by phage display technology." Annu. Rev. Immunol. **12**: 433-55.

CLAIMS:

1. A DNA construct consisting of a transcription unit useful for covalent intracellular joining of selected exons from transcripts of two different genes A and B in a cell in which said genes A and B are expressible, said transcription unit comprising promoter/enhancer elements and a template for RNA synthesis, wherein said template encodes an RNA transcript of the general formula:

RCA - SP - RCB

wherein

RCA represents a nucleotide segment having one or more sequences, each reverse-complementary to one or more sequences of pre-mRNA of gene A and genes related thereto, said sequences of pre-mRNA of gene A and genes related thereto being situated downstream to an exon selected to be spliced to pre-mRNA of gene B or genes related thereto;

SP is either a spacer sequence or a sequence of the formula:

SP1 - LEX - SP2

wherein

SP1 represents a spacer sequence;

LEX represents an exon, said exon encoding a flexible peptide linker or a part thereof preceded by branch point and acceptor splice sequences and followed by a donor splice sequence; and

SP2 represents a spacer sequence; and

RCB represents a nucleotide segment having one or more sequences, each reverse-complementary to one or more sequences of pre-mRNA of gene B and genes related thereto, said sequences of pre-mRNA of gene B and genes related thereto being situated upstream to an exon in pre-mRNA of gene B or genes related thereto selected to be spliced to said selected exon of gene A.

2. A DNA construct of claim 1 wherein said cell in which genes A and B are expressible is an immune cell.

3. A DNA construct of claim 2 wherein said immune cell is a B lymphocyte and either said gene A encodes an antibody light (L) chain and said gene B encodes an antibody heavy (H) chain, or said gene A encodes an antibody heavy (H) chain and said gene B encodes an antibody light (L) chain.

4. A DNA construct of claim 3 wherein said antibody is a mammalian antibody.

5. A DNA construct of claim 4 wherein said mammalian antibody is a mouse or human antibody.

6. A DNA construct of any one of claims 1 to 5 for construction of a mouse antibody library, wherein:

RCA represents a nucleotide segment of about 120 nucleotides, having one sequence reverse-complementary to the 3' region of the J-C intron of the mouse κ chain gene, said 3' region having the sequence

CTTATCTGTA GGGATAAGCG TGCTTTTTTG TGTGTTGTAT
ATAACATAAC TGTTTACACA TAATACACTG AAATGGAGCC
CTTCCTTGTT ACTTCATACC ATCCTCTGTG CTTCCTTCCT C

SP1 represents a spacer sequence;

LEX is a segment encoding a linker peptide used for construction of scFv molecules or a part of said peptide;

SP2 represents a spacer sequence; and

RCB represents a nucleotide segment containing 19 different sequences, each of which is reverse-complementary to a defined stretch of about 20 nucleotides which is centered in the leader exon/intron junction of one or more germ-line mouse VH genes.

7. A DNA construct of claim 6, wherein the 19 sequences contained in RCB are reverse-complementary to the sequences 1 to 19 depicted in Fig.9.

8. A DNA construct of claim 7, wherein the segment RCB is encoded by the sequence depicted in Fig. 11.

9. A DNA construct of any one of claims 6 to 8 having the schematic structure depicted in Fig. 12, wherein RCA, SP1, LEX, SP2 and RCB correspond to the stretches Xh-N, Cgamma1-hinge intron, linker, mJH3-JH4 intron and Xb-R, respectively.

10. A DNA construct of any one of claims 1 to 5 for construction of a human antibody library, wherein:

RCA represents a nucleotide segment having the following three sequences: (a) a sequence of about 120 nucleotides reverse complementary to the 3' end of the J-C intron of the human Ig κ chain gene; (b) a sequence of about 120 nucleotides reverse complementary to the 3' end of the Jλ1-Cλ1 intron of the human Ig λ1 gene; and (c) a sequence of about 50 nucleotides reverse complementary to the 3' end of the Jλ2-Cλ2 intron of the human Ig λ2 gene;

SP1 represents a spacer sequence;

LEX is a segment encoding a linker peptide used for construction of scFv molecules or a part of said peptide;

SP2 represents a spacer sequence; and

RCB represents a nucleotide segment containing 24 different sequences, each of which is reverse-complementary to a defined stretch of about 20-25 nucleotides which is centered in the leader exon/intron junction of one or more germ-line mouse VH genes.

11. A DNA construct of claim 10, wherein the 24 sequences contained in RCB are reverse-complementary to the sequences 1 to 24 depicted in Fig. 13.

12. A DNA construct of claim 11, wherein the segment RCB is encoded by the sequence depicted in Fig. 14.

13. A DNA construct of any one of claims 10 to 12, wherein the segment RCA is encoded by the sequence depicted in Fig. 16.

14. A DNA construct of any one of claims 10 to 13 having the schematic structure depicted in Fig. 17A wherein RCA, SP1, LEX, SP2 and RCB correspond to the stretches Xh-N, Cgamma1-hinge intron, linker, mJH3-JH4 intron and Xb-R, respectively.

15. A transgenic non-human vertebrate harboring a DNA construct of any one of claims 1 to 14.

16. A transgenic non-human vertebrate of claim 15 wherein said vertebrate is a mouse.

17. A transgenic mouse of claim 16, wherein said mouse is a mouse of ordinary mouse strain.

18. A transgenic mouse of claim 16, wherein said mouse contains human Ig H and L chain gene loci and is incapable of producing self-antibodies.

19. A method for generating a variegated population of cDNA molecules suitable for preparation of gene libraries encoding scFv molecules of antibodies of interest, said method comprising the steps:

(a) immunizing a transgenic mouse of claim 17 or 18 with an appropriate antigen to produce the antibodies of interest in said transgenic mouse;

(b) extracting RNA from mature B cells of said immunized mouse; and

(c) subjecting said RNA preparation to reverse transcriptase reaction,

whereby a variegated population of cDNA molecules are obtained from which gene libraries encoding scFv molecules of said antibodies of interest can be produced by PCR employing variable region-specific primers.

20. A library comprising a variegated population of cDNA molecules suitable for preparation of gene libraries encoding scFv molecules of antibodies of interest, produced by the method of claim 19.

21. A method for generating a variegated population of double-stranded DNA molecules encoding scFv molecules of antibodies of interest, said method comprising the steps:

(a) immunizing a transgenic mouse of claim 17 or 18 with an appropriate antigen to produce the antibodies of interest in said transgenic mouse;

(b) extracting RNA from mature B cells of said immunized mouse; and

(c) subjecting said RNA preparation to reverse transcriptase-PCR,

thus obtaining a variegated population of double-stranded DNA molecules encoding scFv molecules of said antibodies of interest.

22. A gene library of variegated double-stranded DNA molecules encoding scFv molecules of antibodies of interest, produced by the method of claim 21.

23. A gene library of claim 22 wherein said DNA molecules encode scFv molecules of mouse antibodies.

24. A gene library of claim 22 wherein said DNA molecules encode scFv molecules of human antibodies.

25. An expression vector comprising a variegated population of DNA molecules encoding scFv molecules of antibodies of interest of any one of claims 22-24.

26. An expression vector of claim 25 which is expressible and displayable on the surface of a cell or viral particle.

27. An expression vector of claim 25 or 26 which is a phagemid.

28. A phage-display library comprising a plurality of recombinant phages each having an expression vector of any one of claims 25 to 27.

29. A phage-display library of claim 28 wherein said expression vector comprises DNA molecules encoding scFv molecules of antibodies of interest fused to a capsid protein of the phage.

30. A phage-display library of claim 29 wherein said antibodies of interest are mouse or human antibodies.

31. A method for generating a phage-display library of any one of claims 28 to 30 which comprises co-transfecting a suitable bacterial cell with an expression vector of any one of claims 25-27 and a helper phage.

32. A method of claim 31 wherein said bacterial cell is of a E. coli strain and said helper phage is M13 filamentous bacteriophage.

33. A method for generating an antibody library comprising a variegated population of scFv molecules of antibodies of interest expressed and displayed on the surface of a cell or viral particle, said method comprising the steps:

(a) immunizing a transgenic mouse of claim 17 or 18 with an appropriate antigen to produce the antibodies of interest in said transgenic mouse;

(b) extracting mRNA from mature B cells of said immunized mouse;

(c) subjecting said mRNA preparation to RT-PCR thus obtaining PCR products which consist of a variegated population of DNA molecules encoding scFv molecules of said antibodies of interest;

(d) cloning the PCR products obtained in (c) in an expression vector; and

(e) cloning said expression vector in a suitable system whereby said DNA molecules of (c) encoding scFv molecules of antibodies of interest are expressed and displayed on the surface of a cell or viral particle.

34. A method according to claim 33, wherein said library is a phage-display library and said scFv molecules of antibodies of interest are fused to a capsid protein of the phage.

35. A DNA construct of claim 1 or 2 wherein said immune cell is a human T lymphocyte and either said gene A encodes a T-cell receptor (TCR) α chain and said gene B encodes a TCR β chain, or said gene A encodes a TCR β chain and said gene B encodes a TCR α chain.

36. A method for generating a variegated population of cDNA molecules suitable for preparation of gene libraries encoding scFv molecules of human TCRs of interest, said method comprising the steps:

(a) introducing a DNA construct of claim 35 to T cells isolated from an immune individual;

(b) extracting RNA from the T cells of (a); and

(c) subjecting said RNA preparation to reverse transcriptase reaction,

whereby a variegated population of cDNA molecules are obtained from which gene libraries encoding scFv molecules of said TCRs of interest can be produced by PCR employing variable region-specific primers.

37. A library comprising a variegated population of cDNA molecules suitable for preparation of gene libraries encoding scFv molecules of human TCRs of interest, produced by the method of claim 36.

38. A method for generating a variegated population of double-stranded DNA molecules encoding scFv molecules of human TCRs of interest, said method comprising the steps:

(a) delivering a DNA construct of claim 35 to T cells isolated from an immune individual;

(b) extracting RNA from the T cells of (a); and

(c) subjecting said RNA preparation to reverse transcriptase-PCR,

thus obtaining a variegated population of double-stranded DNA molecules encoding scFv molecules of said human TCRs of interest.

39. A gene library of variegated double-stranded DNA molecules encoding scFv molecules of human TCRs of interest, produced by the method of claim 38.

40. An expression vector comprising a variegated population of DNA molecules encoding scFv molecules of human TCRs of interest of claim 38 or 39.

41. An expression vector of claim 40 which is expressible and displayable on the surface of a cell or viral particle.

42. An expression vector of claim 40 or 41 which is a phagemid.

43. A phage-display library comprising a plurality of recombinant phages each having an expression vector of any one of claims 40 to 42.

44. A phage-display library of claim 43 wherein said expresssion vector comprises DNA molecules encoding scFv molecules of human TCRs fused to a capsid protein of the phage.

45. A method for generating a phage-display library of claim 43 or 44 which comprises co-transfecting a suitable bacterial cell with an expression vector of any one of claims 40 to 42 and a helper phage.

46. A method of claim 45 wherein said bacterial cell is E. coli and said helper phage is M13 filamentous bacteriophage.

47. A method for generating a human TCR library comprising a variegated population of scFv molecules of human TCRs of interest expressed and displayed on the surface of a cell or viral particle, said method comprising the steps:

(a) introducing a DNA construct of claim 35 to T cells isolated from an immune individual;

(b) extracting RNA from the T cells of (a);

(c) subjecting said mRNA preparation to RT-PCR, thus obtaining PCR products which consist of a variegated population of DNA molecules encoding scFv molecules of said human TCRs of interest;

(d) cloning the PCR products obtained in (c) in an expression vector; and

(e) cloning said expression vector in a suitable system whereby said DNA molecules of (c) encoding scFv molecules of human TCRs of interest are expressed and displayed on the surface of a cell or viral particle.

48. A method according to claim 47, wherein said library is a phage-display library and said scFv molecules of human TCRs of interest are fused to a capsid protein of the phage.

Figure 1

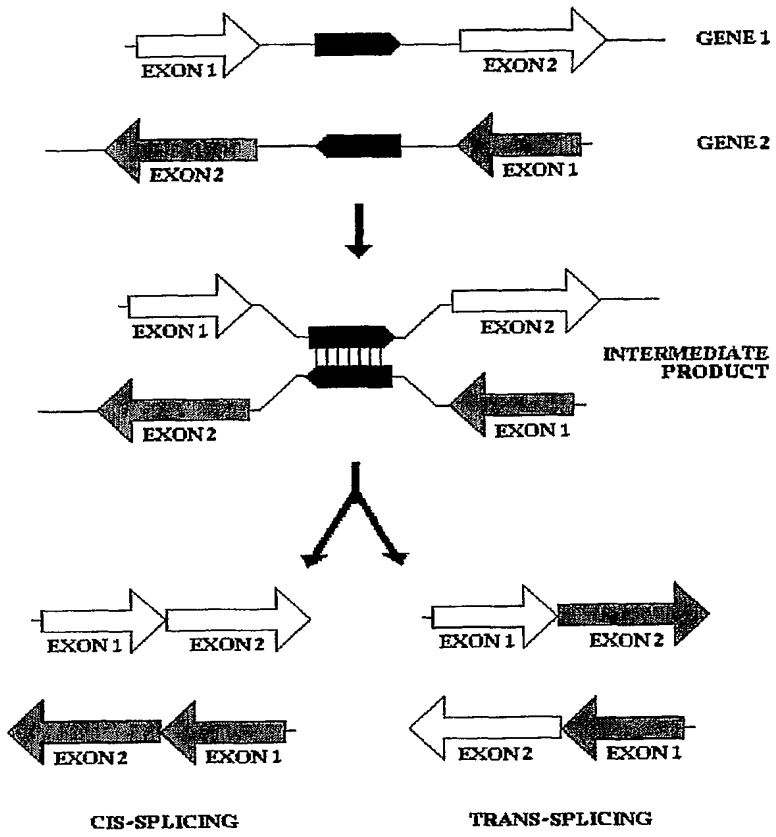

GENE 1
EXON 1
EXON 2
GENE 2
EXON 2
EXON 1
EXON 1
EXON 2
INTERMEDIATE PRODUCT
EXON 2
EXON 1
EXON 1
EXON 2
EXON 1
EXON 2
EXON 2
EXON 1
CIS-SPLICING
TRANS-SPLICING

Figure 2

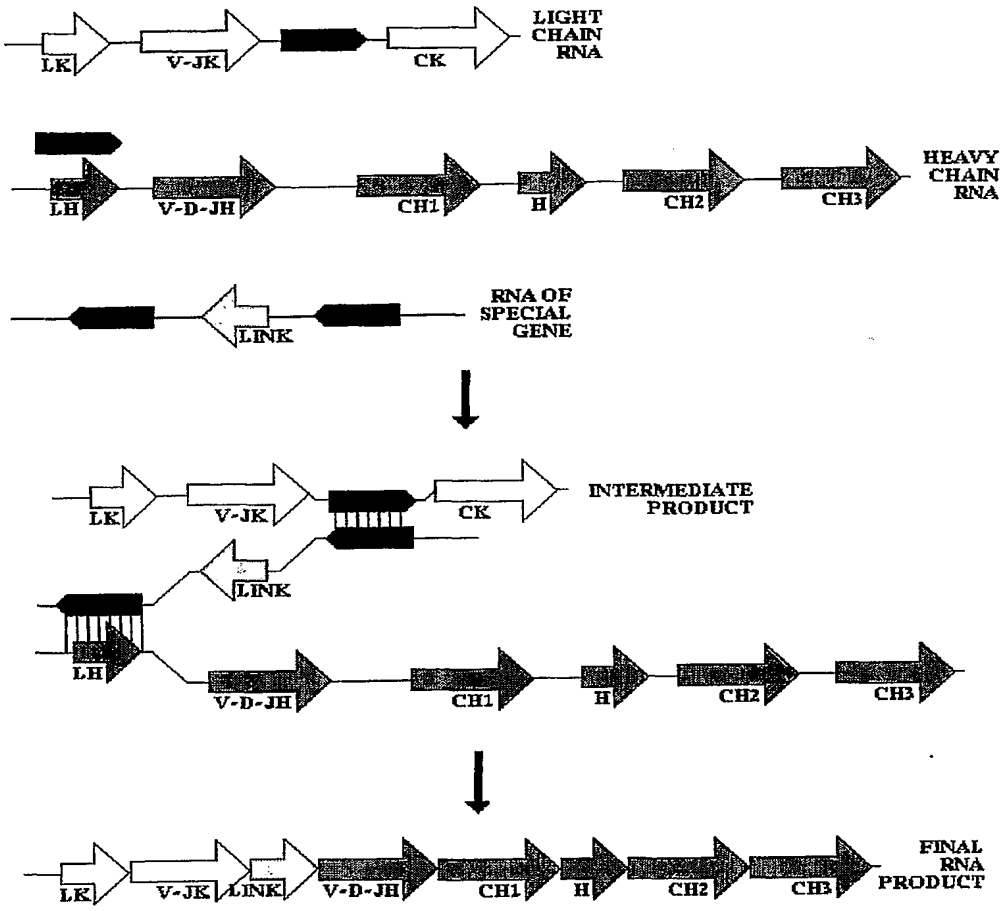

LK
V-JK
CK
LIGHT CHAIN RNA
LH
V-D-JH
CH1
H
CH2
CH3
HEAVY CHAIN RNA
LINK
RNA OF SPECIAL GENE
INTERMEDIATE PRODUCT
FINAL RNA PRODUCT

Figure 3

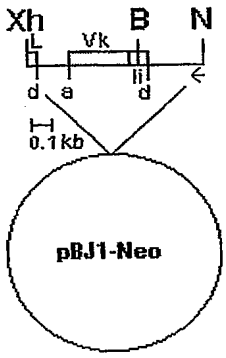


Figure 4

Linker ... CCC GGA TCC GGT GAA GGA G/gtaagt
ctttctctccatag / GT GTC CAT TGC CAG...VH
Linker ... CCC GGA TCC GGT GAA GGA GGT GTC CAT TGC CAG...VH

Figure 5

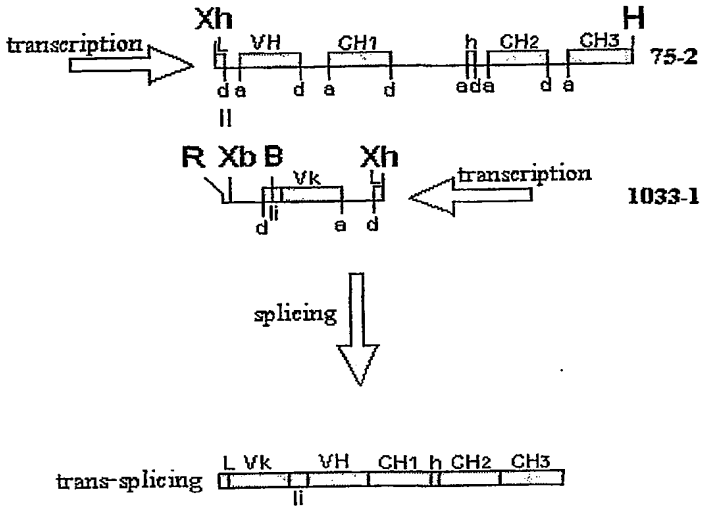


Figure 6

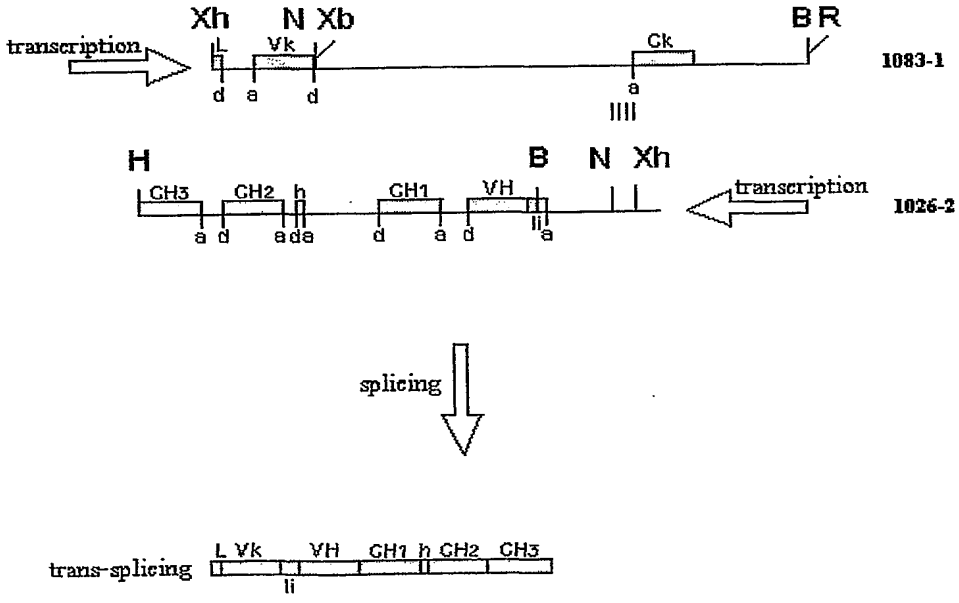
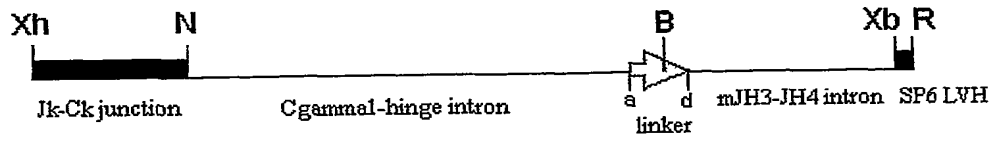


Figure 7

A.



B.

```
 Xho I
CTCGAGGAAG GAAGCACAGA GGATGGTATG AAGTAACAAG GAAGGGCTCC  50
GAGCTCCTTC CTTCGTGTCT CCTACCATAC TTCATTGTTC CTTCCCGAGG

ATTTCAGTGT ATTATGTGTA AACAGTTATG TTATATACAG ACACACAAAA 100
TAAAGTCACA TAATACACAT TTGTCAATAC AATATATGTC TGTGTGTTTT
                              Not I
AAGCACGCTT ATCCCTACAG ATAAGCGGCC GCCAGCACAG GGAGGGAGGG 150
TTCGTGCGAA TAGGGATGTC TATTCGCCGG CGGTCGTGTC CCTCCCTCCC

TGTCTGCTGG AAGCCAGGCT CAGCGCTCCT GCCTGGACGC ATCCCGGCTA 200
ACAGACGACC TTCGGTCCGA GTCGCGAGGA CGGACCTGCG TAGGGCCGAT

TGCAGCCCCA GTCCAGGGCA GCAAGGCAGG CCCCGTCTGC CTCTTCACCC 250
ACGTCGGGGT CAGGTCCCGT CGTTCCGTCC GGGGCAGACG GAGAAGTGGG

GGAGGCCTCT GCCCGCCCCA CTCATGCTCA GGGAGAGGGT CTTCTGGCTT 300
CCTCCGGAGA CGGGCGGGGT GAGTACGAGT CCCTCTCCCA GAAGACCGAA

TTTCCCCAGG CTCTGGGCAG GCACAGGCTA GGTGCCCCTA ACCCAGGCCC 350
AAAGGGGTCC GAGACCCGTC CGTGTCCGAT CCACGGGGAT TGGGTCCGGG

TGCACACAAA GGGGCAGGTG CTGGGCTCAG ACCTGCCAAG AGCCATATCC 400
ACGTGTGTTT CCCCGTCCAC GACCCGAGTC TGGACGGTTC TCGGTATAGG

GGGAGGACCC TGCCCCTGAC CTAAGCCCAC CCCAAAGGCC AAACTCTCCA 450
CCCTCCTGGG ACGGGGACTG GATTCGGGTG GGGTTTCCGG TTTGAGAGGT

CTCCCTCAGC TCGGACACCT TCTCTCCTCC CAGATTCCAG TAACTCCCAA 500
GAGGGAGTCG AGCCTGTGGA AGAGAGGAGG GTCTAAGGTC ATTGAGGGTT
                                    BamH I
TCTTCTCTCT GCAGAATCTA CTTCCGGTTC AGGAAAGCCC GGATCCGGTG 550
AGAAGAGAGA CGTCTTAGAT GAAGGCCAAG TCCTTTCGGG CCTAGGCCAC

AAGGAGGTAA GTTGCACAGG CAGGGAACAG AATGTGGAAC AATGACTTGA 600
TTCCTCCATT CAACGTGTCC GTCCCTTGTC TTACACCTTG TTACTGAACT

ATGGTTGATT CTTGTGTGAC ACCAGGAATT GGCATAATGT CTGAGTTGCC 650
TACCAACTAA GAACACACTG TGGTCCTTAA CCGTATTACA GACTCAACGG
```

```
CAGGGGTGAT TCTAGTCAGA CTCTGGGGTT TTTGTCGGGT ACAGAGGAAA 700
GTCCCCACTA AGATCAGTCT GAGACCCCAA AAACAGCCCA TGTCTCCTTT
              Xba I                        EcoR I
AACCCACTAT TGTCTAGAGA GCCCCTTACC TGCAATTAGA ATTC
TTGGGTGATA ACAGATCTCT CGGGGAATGG ACGTTAATCT TAAG
```

Figure 8

A.

Linker ... CCC GGA TCC GGT GAA GGA G/gtaagt
ctttctctccatag / GT GTC CAT TGC CAG...VH
Linker ... CCC GGA TCC GGT GAA GGA GGT GTC CAT TGC CAG...VH

B.

V_K ...AAG CTG GAG CTG AAA C / gtaagt...
tcttctctctgcag / AA TCT ACT TCC GGT TCA ...Linker
V_K ...AAG CTG GAG CTG AAA CAA TCT ACT TCC GGT TCA ...Linker

C.

V_K ... AAG CTG GAG CTG AAA C / gtaagt...
ctttctctccatag / GT GTC CAT TGC CAG ...VH
V_K... AAG CTG GAG CTG AAA CGT GTC CAT TGC CAG ...VH

Figure 9

1. 5' CAGCTACAGGTAAGGGGCTC 3'
   VH-186-2 1 J00530
   C36e 1 L26856
   C31e 1 L26853
   B25c 1 L26880
   C11c 1 L26925
   B12c 1 L26869
   C16c 1 L26928
   VH124 1 K00706
   P2M5 1 D14634
   VH3 1 J00536
   B23c 1 L26879
   B1c 1 L26867
   B26c 1 L26881

2. 5' GCAGCTAACGGTAAGGGGCT 3'
   V23 1 J00534

3. 5' ACAGCTACAGGTAAGGAGCT 3'
   B16e 1 L26874

4. 5' TCAGGAACTGCAGGTAAGGG 3'
   VH105 1 J00507
   BALB71 1 L33959
   BALB17 1 L33954
   C57C27 1 L33934
   C57G9 1 L33947
   BALB11 1 L33952

5. 5' TAACTGCAGGTAAGGGGCTC 3'
   J558-43y 1 M34987
   pM11 1 D14633
   C57G5 1 L33946
   37A11 1 M18951

6. 5' ATCCTGTCAGTAACTACAGGT 3'
   C57C2 1 L33942
   C57G26 1 L33937

7. 5' TAACTACAGGTAAGGGGCTC 3'
   J558-43x 1 M34983

8. 5' TAACTGCAGGTAAGGGGGTC 3'
   M34976 1 M34976

9. 5' TAATTGCAGGTAAGGGGCTC 3'
   SP6VH 1 X56936

10. 5' TTCCCAAGCTGTAAGTGTGTC 3'
    PJ14 2 J00492

11. 5' TTCCCAAGCTGTGAGTGTTTC 3'
    VH101 2 J00502

12. 5' TTTAAAAGGTAATTCATGGAAAA 3'
    V-H-441 4 J00541

13. 5' TTTAAAAGGTAATTTATTGAGAA 3'
    VH10-19 5 X59817
    VH283 5 X00163

14. 5' CTTTTACATGGTAATTTATGGG 3'
    V1 7 J00538

15. 5' ACTTTTAAATGGTAATTTATGGT 3'
    pBV19B4 7 M16725

16. 5' CTGCCCAAAGTAAGACATCAG 3'
    251 9 Z15022
    281 9 Z15023

17. 5' TGGTTATAGGTAAGGGGCTC 3'
    H2b-3 14 M12291

18. 5' TGGTTACAGGTAAGGGGCTC 3'
    V(H)10 14 X03571

19. 5' TGGTTACAGGTAAGGAGCTC 3'
    V(H)4a-3 14 X03572

Figure 10

1. # 8424
5′ - CTA GAG AGC TCC TTA CCT GTA ACC AGA GCC CCT TAC CTG TAA
CCA GAG CCC CTT ACC TAT AAC CAC TGA TGT CTT ACT TTG GGC - 3′

2. # 8425
5′ - P - AGA CCA TAA ATT ACC ATT TAA AAG TGG GCC CCC CAT AAA
TTA CCA TGT AAA AGT TCT CAA TAA ATT ACC TTT TAA ATT
TTC CAT - 3′

3. # 8426
5′ - P - GAA TTA CCT TTT AAA GAA ACA CTC ACA GCT TGG GAA CCG
CGG GAC ACA CTT ACA GCT TGG GAA GAG CCC CTT ACC TGC
AAT TAG -3′

4. # 8427
5′ - P - ACC CCC TTA CCT GCA GTT AGA GCC CCT TAC CTG TAG TTA
ACC TGT AGT TAC TGA CAG GAT ATC GAT GAG CCC CTT ACC
TGC AGT - 3′

5. # 8311
5′ - P - TAC CCT TAC CTG CAG TTC CTG AAG CTC CTT ACC TGT AGC
TGT AGC CCC TTA CCG TTA GCT GCG AGC CCC TTA CCT GTA
GCT GG - 3′

6. # 6727
5′ - AAT TCC AGC TAC AGG TAA GGG GCT CGC AGC TAA CGG TAA GGG
GCT ACA GCT ACA GGT AAG GAG C TT C - 3′

7. # 301230
5′ - P - AGG AAC TGC AGG TAA GGG TAA CTG CAG GTA - 3′

8. # 8428
5′ - P - AGG GGC TCA TCG ATA TCC TGT CAG TAA CTA CAG GTT AAC
TAC AGG TAA GGG GCT CTA ACT GCA GGT AAG GGG GTC TAA
TTG CAG - 3′

9. # 8429
5′ - P - GTA AGG GGC TCT TCC CAA GCT GTA AGT GTG TCC CGC GGT
TCC CAA GCT GTG AGT GTT TCT TTA AAA GGT AAT TCA TGG
AAA ATT - 3′

10. # 8430
5′ - P - TAA AAG GTA ATT TAT TGA GAA CTT TTA CAT GGT AAT TTA
TGG GGG GCC CAC TTT TAA ATG GTA ATT TAT GGT CTG CCC
AAA GTA - 3′

11. # 7026
5′ - P - AGA CAT CAG TGG TTA TAG GTA AGG GGC TCT GGT TAC AGG
TAA GGG GCT CTG GTT ACA GGT AAG GAG CTC T - 3′

Figure 11

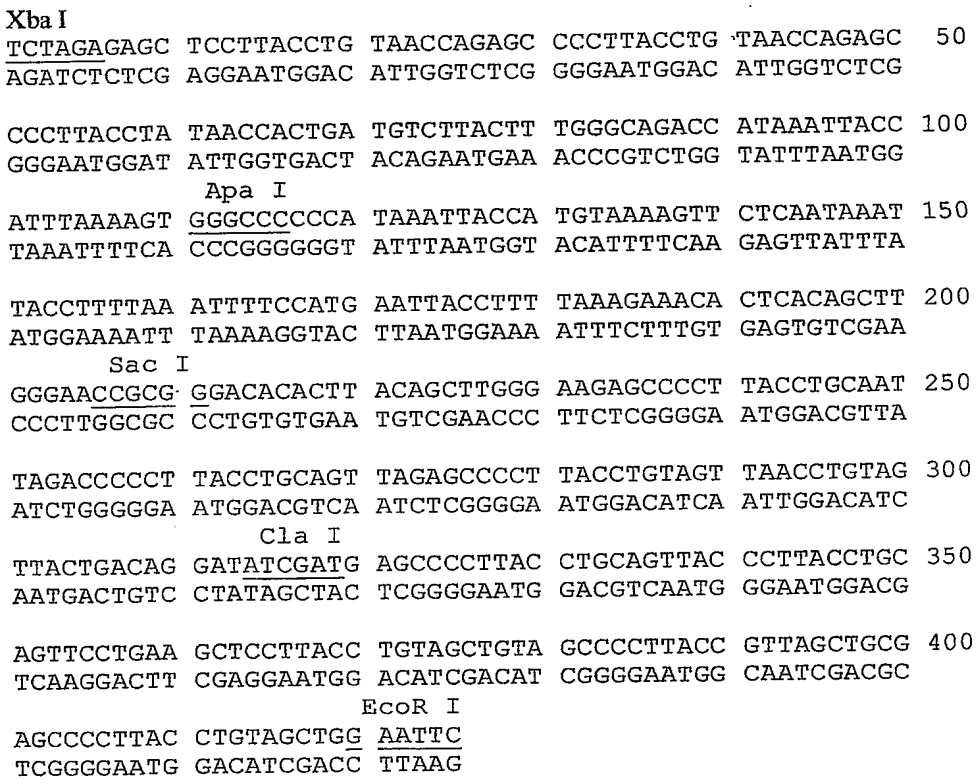
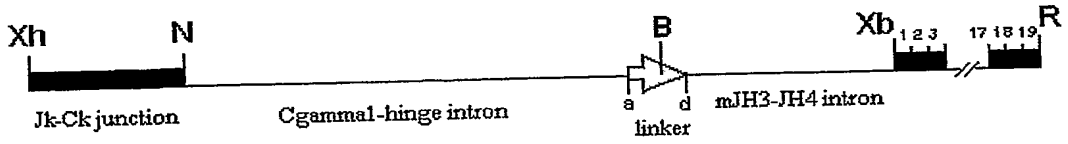

Xba I
TCTAGAGAGC TCCTTACCTG TAACCAGAGC CCCTTACCTG TAACCAGAGC 50
AGATCTCTCG AGGAATGGAC ATTGGTCTCG GGGAATGGAC ATTGGTCTCG
CCCTTACCTA TAACCACTGA TGTCTTACTT TGGGCAGACC ATAAATTACC 100
GGGAATGGAT ATTGGTGACT ACAGAATGAA ACCCGTCTGG TATTTAATGG
Apa I
ATTTAAAAGT GGGCCCCCCA TAAATTACCA TGTAAAAGTT CTCAATAAAT 150
TAAATTTTCA CCCGGGGGGT ATTTAATGGT ACATTTTCAA GAGTTATTTA
TACCTTTTAA ATTTTCCATG AATTACCTTT TAAAGAAACA CTCACAGCTT 200
ATGGAAAATT TAAAAGGTAC TTAATGGAAA ATTTCTTTGT GAGTGTCGAA
Sac I
GGGAACCGCG GGACACACTT ACAGCTTGGG AAGAGCCCCT TACCTGCAAT 250
CCCTTGGCGC CCTGTGTGAA TGTCGAACCC TTCTCGGGGA ATGGACGTTA
TAGACCCCCT TACCTGCAGT TAGAGCCCCT TACCTGTAGT TAACCTGTAG 300
ATCTGGGGGA ATGGACGTCA ATCTCGGGGA ATGGACATCA ATTGGACATC
Cla I
TTACTGACAG GATATCGATG AGCCCCTTAC CTGCAGTTAC CCTTACCTGC 350
AATGACTGTC CTATAGCTAC TCGGGGAATG GACGTCAATG GGAATGGACG
AGTTCCTGAA GCTCCTTACC TGTAGCTGTA GCCCCTTACC GTTAGCTGCG 400
TCAAGGACTT CGAGGAATGG ACATCGACAT CGGGGAATGG CAATCGACGC
EcoR I
AGCCCCTTAC CTGTAGCTGG AATTC
TCGGGGAATG GACATCGACC TTAAG

Figure 12

A.



B.

Xho I
CTCGAGGAAG GAAGCACAGA GGATGGTATG AAGTAACAAG GAAGGGCTCC 50
GAGCTCCTTC CTTCGTGTCT CCTACCATAC TTCATTGTTC CTTCCCGAGG

ATTTCAGTGT ATTATGTGTA AACAGTTATG TTATATACAG ACACACAAAA 100
TAAAGTCACA TAATACACAT TTGTCAATAC AATATATGTC TGTGTGTTTT

Not I
AAGCACGCTT ATCCCTACAG ATAAGCGGCC GCCAGCACAG GGAGGGAGGG 150
TTCGTGCGAA TAGGGATGTC TATTCGCCGG CGGTCGTGTC CCTCCCTCCC

TGTCTGCTGG AAGCCAGGCT CAGCGCTCCT GCCTGGACGC ATCCCGGCTA 200
ACAGACGACC TTCGGTCCGA GTCGCGAGGA CGGACCTGCG TAGGGCCGAT

TGCAGCCCCA GTCCAGGGCA GCAAGGCAGG CCCCGTCTGC CTCTTCACCC 250
ACGTCGGGGT CAGGTCCCGT CGTTCCGTCC GGGGCAGACG GAGAAGTGGG

GGAGGCCTCT GCCCGCCCCA CTCATGCTCA GGGAGAGGGT CTTCTGGCTT 300
CCTCCGGAGA CGGGCGGGGT GAGTACGAGT CCCTCTCCCA GAAGACCGAA

TTTCCCCAGG CTCTGGGCAG GCACAGGCTA GGTGCCCCTA ACCCAGGCCC 350
AAAGGGGTCC GAGACCCGTC CGTGTCCGAT CCACGGGGAT TGGGTCCGGG

TGCACACAAA GGGGCAGGTG CTGGGCTCAG ACCTGCCAAG AGCCATATCC 400
ACGTGTGTTT CCCCGTCCAC GACCCGAGTC TGGACGGTTC TCGGTATAGG

GGGAGGACCC TGCCCCTGAC CTAAGCCCAC CCCAAAGGCC AAACTCTCCA 450
CCCTCCTGGG ACGGGGACTG GATTCGGGTG GGGTTTCCGG TTTGAGAGGT

CTCCCTCAGC TCGGACACCT TCTCTCCTCC CAGATTCCAG TAACTCCCAA 500

```
GAGGGAGTCG AGCCTGTGGA AGAGAGGAGG GTCTAAGGTC ATTGAGGGTT
                                             BamH I
TCTTCTCTCT GCAGAATCTA CTTCCGGTTC AGGAAAGCCC GGATCCGGTG  550
AGAAGAGAGA CGTCTTAGAT GAAGGCCAAG TCCTTTCGGG CCTAGGCCAC

AAGGAGGTAA GTTGCACAGG CAGGGAACAG AATGTGGAAC AATGACTTGA  600
TTCCTCCATT CAACGTGTCC GTCCCTTGTC TTACACCTTG TTACTGAACT

ATGGTTGATT CTTGTGTGAC ACCAGGAATT GGCATAATGT CTGAGTTGCC  650
TACCAACTAA GAACACACTG TGGTCCTTAA CCGTATTACA GACTCAACGG

CAGGGGTGAT TCTAGTCAGA CTCTGGGGTT TTTGTCGGGT ACAGAGGAAA  700
GTCCCCACTA AGATCAGTCT GAGACCCCAA AAACAGCCCA TGTCTCCTTT
             Xba I
AACCCACTAT TGTCTAGAGA GCTCCTTACC TGTAACCAGA GCCCCTTACC  750
TTGGGTGATA ACAGATCTCT CGAGGAATGG ACATTGGTCT CGGGGAATGG

TGTAACCAGA GCCCCTTACC TATAACCACT GATGTCTTAC TTTGGGCAGA  800
ACATTGGTCT CGGGGAATGG ATATTGGTGA CTACAGAATG AAACCCGTCT
                       Apa I
CCATAAATTA CCATTTAAAA GTGGGCCCCC CATAAATTAC CATGTAAAAG  850
GGTATTTAAT GGTAAATTTT CACCCGGGGG GTATTTAATG GTACATTTTC

TTCTCAATAA ATTACCTTTT AAATTTTCCA TGAATTACCT TTTAAAGAAA  900
AAGAGTTATT TAATGGAAAA TTTAAAAGGT ACTTAATGGA AAATTTCTTT
                Sac II
CACTCACAGC TTGGGAACCG CGGGACACAC TTACAGCTTG GGAAGAGCCC  950
GTGAGTGTCG AACCCTTGGC GCCCTGTGTG AATGTCGAAC CCTTCTCGGG

CTTACCTGCA ATTAGACCCC CTTACCTGCA GTTAGAGCCC CTTACCTGTA 1000
GAATGGACGT TAATCTGGGG GAATGGACGT CAATCTCGGG GAATGGACAT
                          Cla I
GTTAACCTGT AGTTACTGAC AGGATATCGA TGAGCCCCTT ACCTGCAGTT 1050
CAATTGGACA TCAATGACTG TCCTATAGCT ACTCGGGGAA TGGACGTCAA

ACCCTTACCT GCAGTTCCTG AAGCTCCTTA CCTGTAGCTG TAGCCCCTTA 1100
TGGGAATGGA CGTCAAGGAC TTCGAGGAAT GGACATCGAC ATCGGGGAAT
                                   EcoR I
CCGTTAGCTG CGAGCCCCTT ACCTGTAGCT GGAATTC
GGCAATCGAC GCTCGGGGAA TGGACATCGA CCTTAAG
```

Figure 13

1. 5' CAGCCACAGGTAAGAGGCTC 3'
   VH1-2 X07448

2. 5' GCAGCCACAGGTAAGGGGCT 3'
   VH1-3 X62109
   VH1-45 X92209

3. 5' CAGCTACAAGTAAGGGGCTT 3'
   VH1-8 M99637
   VH1-69 L22582

4. 5' CACCAACAGGTAACGGACTC 3'
   VH1-18 M99641

5. 5' CAGGCAAGAGAATCCTGAGTT 3'
   VH1-24 M99642

6. 5' TAGCTCCAGGTAAAGGGCCA 3'
   VH1-46 X92343

7. 5' CAGCGACAGGCAAGGAGATG 3'
   VH1-58 M29809

8. 5' TCCCTTCATGTGAGTGCTGT 3'
   VH2-5 X62111

9. 5' CCCCTTCCTGTGAGTGCTGT 3'
   VH2-26 M99648

10. 5' TCCCGTCCTGTGAGTGCTGT 3'
    VH2-70 L21969

11. 5' ATTTTAGAAGGTGATTCATGGA 3'
    VH3-7 M99649
    VH3-48 M99675

12. 5' TAAAAGGTGATTTATGGAGAAC 3'
    VH3-11 M99652
    VH3-15 X92216

13. 5' ATATTAGAAGGTGATTCATGGA 3'
    VH3-13 X92217

14. 5' TATTTTAAAAGGTGATTCATGG 3'
    VH3-9 M99651
    VH3-20 M99657
    VH3-43 M99672
    VH3-53 M99679
    VH3-66 X99218
    VH3-73 L15467
    VH3-74 J00239

15. 5' ATTTTAGAAGGTGAATATGGA 3'
    VH3-21 M99658

16. 5' AAATAAAAGGTAATTCATGGAG 3'
    VH3-23 M99660

17. 5' GCTCTTTTAAGAGGTGATTCAT 3'
    VH3-30 M83134
    VH3-33 M99665

18. 5' TATTTTAAAAGGTAATTCATGGT 3'
    VH3-49 M99676

19. 5' ATTTTTAAAGGTGATTCATGAG 3'
    VH3-64 M99682

20. 5' TATTTTACAAGGTGATTTATGGA 3'
    VH3-72 X92206

21. 5' GGCAGCTCCCAGATGTGAGT 3'
    VH4-4 X62112
    VH4-31 M99683
    VH4-34 M99684
    VH4-59 M29812
    VH4-61 M29811

22. 5' TGGTGGCGGCTCCCAGATGTGAAT 3'
    VH4-39 X05715

23. 5' TTCTCCAAGGTCAGTCCTGC 3'
    VH5-51 M99686

24. 5' CCTCCCATGGGGTCAGTGTC 3'
    VH6-1 X92224

Figure 14

```
 Xba I
TCTAGAGACA CTGACCCCAT GGGAGGGCAG GACTGACCTT GGAGAAATTC  50
AGATCTCTGT GACTGGGGTA CCCTCCCGTC CTGACTGGAA CCTCTTTAAG

ACATCTGGGA GCCGCCACCA ACTCACATCT GGGAGCTGCC TCCATAAATC 100
TGTAGACCCT CGGCGGTGGT TGAGTGTAGA CCCTCGACGG AGGTATTTAG
              Apa I
ACCTTGTAAA ATAGGGCCCC TCATGAATCA CCTTTAAAAA TACCATGAAT 150
TGGAACATTT TATCCCGGGG AGTACTTAGT GGAAATTTTT ATGGTACTTA

TACCTTTTAA AATAATGAAT CACCTCTTAA AAGAGCCTCC ATGAATTACC 200
ATGGAAAATT TTATTACTTA GTGGAGAATT TTCTCGGAGG TACTTAATGG
                                Hind III
TTTTATTTTC CATATTCACC TTCTAAAATA AGCTTCCATG AATCACCTTT 250
AAAATAAAAG GTATAAGTGG AAGATTTTAT TCGAAGGTAC TTAGTGGAAA

TAAAATATCC ATGAATCACC TTCTAATATG TTCTCCATAA ATCACCTTTT 300
ATTTTATAGG TACTTAGTGG AAGATTATAC AAGAGGTATT TAGTGGAAAA
                                                 Cla I
ATCCATGAAT CACCTTCTAA AATACAGCAC TCACAGGACG GGAATCGATA 350
TAGGTACTTA GTGGAAGATT TTATGTCGTG AGTGTCCTGC CCTTAGCTAT

CAGCACTCAC AGGAAGGGGA CAGCACTCAC ATGAAGGGAC ATCTCCTTGC 400
GTCGTGAGTG TCCTTCCCCT GTCGTGAGTG TACTTCCCTG TAGAGGAACG

CTGTCGCTGT GGCCCTTTAC CTGGAGCTAA ACTCAGGATT CTCTTGCCTG 450
GACAGCGACA CCGGGAAATG GACCTCGATT TGAGTCCTAA GAGAACGGAC
Sac II
CCGCGGGAGT CCGTTACCTG TTGGTGAAGC CCCTTACTTG TAGCTGAGCC 500
GGCGCCCTCA GGCAATGGAC AACCACTTCG GGGAATGAAC ATCGACTCGG
                                      EcoR I
CCTTACCTGT GGCTGCGAGC CTCTTACCTG TGGCTGGAAT TC
GGAATGGACA CCGACGCTCG GAGAATGGAC ACCGACCTTA AG
```

Figure 15

1. (78 mer)
5′ – TCG AGC TGC GGG GTG GAT GAG GGG CAG GGG GTC AGA GTC CTG GTG
TCC ACC TGG GCC GCG GCT GTG TGG ACA GGG AAG – 3′

2. (81 mer)
5′ – P – GGG GTG AGA GAG GGC AGA CAG AAT ACC GGG GTG TTG TGG AGC
CCC TCT CTC TGT CTA AAC TCT CTG GGA GGG TTC ACA GTG – 3′

3. (73 mer)
5′ – P – TGG CCA TCC GGT CCA CCC GGG GTC TAG ACT GAG GAA AGA AGC
AAA CAG GAT GGT GTT TAA GTA ACA AAG TTC T – 3′

4. (77 mer)
5′ – P – GCC CTT GGG TGT GTT GTT TGC GGA TAA TCA CAG GGC ATG TTA
GGG ACA GAC AGA AAA CAG CAT GCT TAT CCC AGA GC – 3′

5. (61 mer)
5′ – GGC CGC TCT GGG ATA AGC ATG CTG TTT TCT GTC TGT CCC TAA CAT
GCC CTG TGA TTA TCC G – 3′

6. (30 mer)
5′ – P – CAA ACA ACA CAC CCA AGG GCA GAA CTT TGT – 3′

7. (73 mer)
5′ – P – TAC TTA AAC ACC ATC CTG TTT GCT TCT TTC CTC AGT CTA
GAC CCC GGG TGG ACC GGA TGG CCA CAC TGT GAA C – 3′

8. (81 mer)
5′ – P – CCT CCC AGA GAG TTT AGA CAG AGA GAG GGG CTC CAC AAC ACC
CCG GTA TTC TGT CTG CCC TCT CTC ACC CCC TTC CCT GTC – 3′

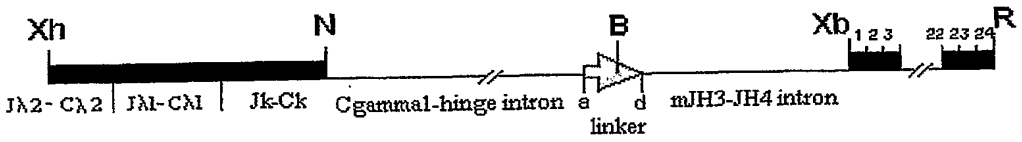
9. (64 mer)
5′ – P – CAC ACA GCC GCG GCC CAG GTG GAC ACC AGG ACT CTG ACC CCC
TGC CCC TCA TCC ACC CCG CAG C – 3′

Figure 16

```
Xho I
CTCGAGCTGC GGGGTGGATG AGGGGCAGGG GGTCAGAGTC CTGGTGTCCA  50
GAGCTCGACG CCCCACCTAC TCCCCGTCCC CCAGTCTCAG GACCACAGGT
      Sac II
CCTGGGCCGC GGCTGTGTGG ACAGGGAAGG GGGTGAGAGA GGGCAGACAG 100
GGACCCGGCG CCGACACACC TGTCCCTTCC CCCACTCTCT CCCGTCTGTC

AATACCGGGG TGTTGTGGAG CCCCTCTCTC TGTCTAAACT CTCTGGGAGG 150
TTATGGCCCC ACAACACCTC GGGGAGAGAG ACAGATTTGA GAGACCCTCC
                                  Xba I
GTTCACAGTG TGGCCATCCG GTCCACCCGG GGTCTAGACT GAGGAAAGAA 200
CAAGTGTCAC ACCGGTAGGC CAGGTGGGCC CCAGATCTGA CTCCTTTCTT

GCAAACAGGA TGGTGTTTAA GTAACAAAGT TCTGCCCTTG GGTGTGTTGT 250
CGTTTGTCCT ACCACAAATT CATTGTTTCA AGACGGGAAC CCACACAACA

TTGCGGATAA TCACAGGGCA TGTTAGGGAC AGACAGAAAA CAGCATGCTT 300
AACGCCTATT AGTGTCCCGT ACAATCCCTG TCTGTCTTTT GTCGTACGAA
         Not I
ATCCCAGAGC GGCCGC
TAGGGTCTCG CCGGCG
```

Figure 17

A.

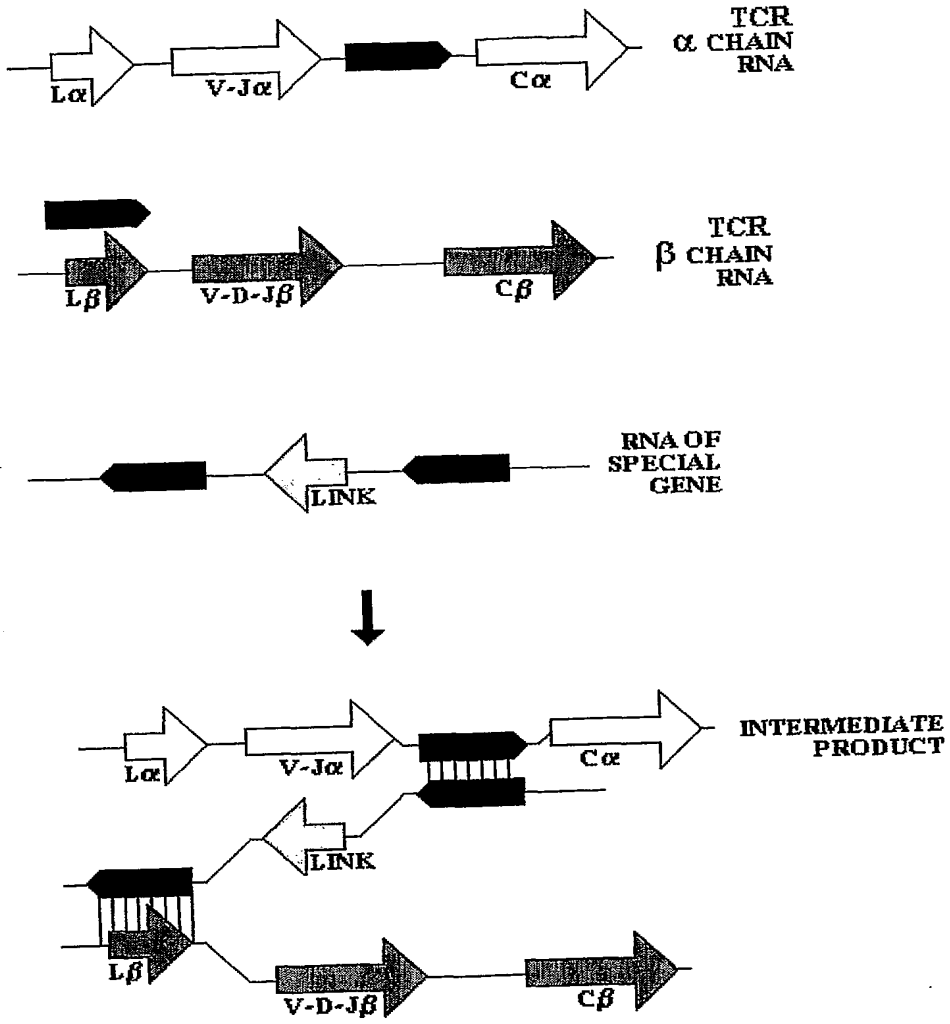


B.

```
 Xho I
CTCGAGCTGC GGGGTGGATG AGGGGCAGGG GGTCAGAGTC CTGGTGTCCA    50
GAGCTCGACG CCCCACCTAC TCCCCGTCCC CCAGTCTCAG GACCACAGGT
      Sac II
CCTGGGCCGC GGCTGTGTGG ACAGGGAAGG GGGTGAGAGA GGGCAGACAG   100
GGACCCGGCG CCGACACACC TGTCCCTTCC CCCACTCTCT CCCGTCTGTC

AATACCGGGG TGTTGTGGAG CCCCTCTCTC TGTCTAAACT CTCTGGGAGG   150
TTATGGCCCC ACAACACCTC GGGGAGAGAG ACAGATTTGA GAGACCCTCC
                                      Xba I
GTTCACAGTG TGGCCATCCG GTCCACCCGG GGTCTAGACT GAGGAAAGAA   200
CAAGTGTCAC ACCGGTAGGC CAGGTGGGCC CCAGATCTGA CTCCTTTCTT

GCAAACAGGA TGGTGTTTAA GTAACAAAGT TCTGCCCTTG GGTGTGTTGT   250
CGTTTGTCCT ACCACAAATT CATTGTTTCA AGACGGGAAC CCACACAACA

TTGCGGATAA TCACAGGGCA TGTTAGGGAC AGACAGAAAA CAGCATGCTT   300
AACGCCTATT AGTGTCCCGT ACAATCCCTG TCTGTCTTTT GTCGTACGAA
          Not I
ATCCCAGAGC GGCCGCCAGC ACAGGGAGGG AGGGTGTCTG CTGGAAGCCA   350
TAGGGTCTCG CCGGCGGTCG TGTCCCTCCC TCCCACAGAC GACCTTCGGT

GGCTCAGCGC TCCTGCCTGG ACGCATCCCG GCTATGCAGC CCCAGTCCAG   400
CCGAGTCGCG AGGACGGACC TGCGTAGGGC CGATACGTCG GGGTCAGGTC

GGCAGCAAGG CAGGCCCCGT CTGCCTCTTC ACCCGGAGGC CTCTGCCCGC   450
CCGTCGTTCC GTCCGGGGCA GACGGAGAAG TGGGCCTCCG GAGACGGGCG

CCCACTCATG CTCAGGGAGA GGGTCTTCTG GCTTTTTCCC CAGGCTCTGG   500
GGGTGAGTAC GAGTCCCTCT CCCAGAAGAC CGAAAAAGGG GTCCGAGACC

GCAGGCACAG GCTAGGTGCC CCTAACCCAG GCCCTGCACA CAAAGGGGCA   550
CGTCCGTGTC CGATCCACGG GGATTGGGTC CGGGACGTGT GTTTCCCCGT

GGTGCTGGGC TCAGACCTGC CAAGAGCCAT ATCCGGGAGG ACCCTGCCCC   600
CCACGACCCG AGTCTGGACG GTTCTCGGTA TAGGCCCTCC TGGGACGGGG

TGACCTAAGC CCACCCCAAA GGCCAAACTC TCCACTCCCT CAGCTCGGAC   650
ACTGGATTCG GGTGGGGTTT CCGGTTTGAG AGGTGAGGGA GTCGAGCCTG

ACCTTCTCTC CTCCCAGATT CCAGTAACTC CCAATCTTCT CTCTGCAGAA   700
TGGAAGAGAG GAGGGTCTAA GGTCATTGAG GGTTAGAAGA GAGACGTCTT
                          BamH I
TCTACTTCCG GTTCAGGAAA GCCCGGATCC GGTGAAGGAG GTAAGTTGCA   750
AGATGAAGGC CAAGTCCTTT CGGGCCTAGG CCACTTCCTC CATTCAACGT

CAGGCAGGGA ACAGAATGTG GAACAATGAC TTGAATGGTT GATTCTTGTG   800
GTCCGTCCCT TGTCTTACAC CTTGTTACTG AACTTACCAA CTAAGAACAC
```

```
TGACACCAGG AATTGGCATA ATGTCTGAGT TGCCCAGGGG TGATTCTAGT  850
ACTGTGGTCC TTAACCGTAT TACAGACTCA ACGGGTCCCC ACTAAGATCA
                                                Xba I
CAGACTCTGG GGTTTTTGTC GGGTACAGAG GAAAAACCCA CTATTGTCTA  900
GTCTGAGACC CCAAAAACAG CCCATGTCTC CTTTTTGGGT GATAACAGAT

GAGAGCCTCT TACCTGTGGC TGAGCCCCTT ACCTGTGGCT GCAAGCCCCT  950
CTCTCGGAGA ATGGACACCG ACTCGGGGAA TGGACACCGA CGTTCGGGGA

TACTTGTAGC TGGAGTCCGT TACCTGTTGG TGAACTCAGG ATTCTCTTGC 1000
ATGAACATCG ACCTCAGGCA ATGGACAACC ACTTGAGTCC TAAGAGAACG
   Apa I
CTGGGGCCCT GGCCCTTTAC CTGGAGCTAC ATCTCCTTGC CTGTCGCTGA 1050
GACCCCGGGA CCGGGAAATG GACCTCGATG TAGAGGAACG GACAGCGACT

CAGCACTCAC ATGAAGGGAA CAGCACTCAC AGGAAGGGGA CAGCACTCAC 1100
GTCGTGAGTG TACTTCCCTT GTCGTGAGTG TCCTTCCCCT GTCGTGAGTG
          Hind III
AGGACGGGAA AGCTTTCCAT GAATCACCTT CTAAAATGTT CTCCATAAAT 1150
TCCTGCCCTT TCGAAAGGTA CTTAGTGGAA GATTTTACAA GAGGTATTTA

CACCTTTTAT CCATGAATCA CCTTCTAATA TCCATGAATC ACCTTTTAAA 1200
GTGGAAAATA GGTACTTAGT GGAAGATTAT AGGTACTTAG TGGAAAATTT
                              Cla I
ATATCCATAT TCACCTTCTA AAATATCGAT CTCCATGAAT TACCTTTTAT 1250
TATAGGTATA AGTGGAAGAT TTTATAGCTA GAGGTACTTA ATGGAAAATA

TTATGAATCA CCTCTTAAAA GAGCACCATG AATTACCTTT TAAAATACTC 1300
AATACTTAGT GGAGAATTTT CTCGTGGTAC TTAATGGAAA ATTTTATGAG
                                          Sac II
ATGAATCACC TTTAAAAATT CCATAAATCA CCTTGTAAAA TACCGCGGAC 1350
TACTTAGTGG AAATTTTTAA GGTATTTAGT GGAACATTTT ATGGCGCCTG

TCACATCTGG GAGCTGCCAT TCACATCTGG GAGCCGCCAC CAGCAGGACT 1400
AGTGTAGACC CTCGACGGTA AGTGTAGACC CTCGGCGGTG GTCGTCCTGA
                                    EcoR I
GACCTTGGAG AAGACACTGA CCCCATGGGA GGGAATTC
CTGGAACCTC TTCTGTGACT GGGGTACCCT CCCTTAAG
```

Figure 18


TCR
α CHAIN
RNA
Lα
V-Jα
Cα
TCR
β CHAIN
RNA
Lβ
V-D-Jβ
Cβ
RNA OF
SPECIAL
GENE
LINK
INTERMEDIATE
PRODUCT
Lα
V-Jα
Cα
LINK
Lβ
V-D-Jβ
Cβ